The EMBO Journal vol.5 no.12 pp.3185–3193, 1986

Upstream sequences required for tissue-specific activation of the cardiac actin gene in *Xenopus laevis* embryos

T.J.Mohun, N.Garrett and J.B.Gurdon

C.R.C. Molecular Embryology Group, Department of Zoology, Downing Street, Cambridge CB2 3EJ, UK

Communicated by J.B.Gurdon

The entire DNA sequence of the *Xenopus laevis* cardiac actin gene was determined. A recombinant plasmid comprising the cardiac actin gene promoter fused to the bacterial chloramphenicol acetyl transferase (CAT) gene is correctly regulated after introduction into fertilized *Xenopus* eggs. The fusion gene shows a temporal and tissue-specific pattern of expression in the early embryo which is indistinguishable from that of the endogenous cardiac actin gene. The fusion gene is also activated in cultured embryo fragments that are induced by cell interactions to form embryonic muscle tissue. Tissue-specific expression of the recombinant requires sequences between 217 and 416 nucleotides upstream from the transcription initiation site. In contrast, both the chimaeric gene and the entire cardiac actin gene are expressed at a basal level after microinjection into *Xenopus* oocytes, requiring only the presence of a TATA box upstream from the cap site.

Key words: actin genes/Amphibia/gene transfer/transcriptional control/upstream DNA sequence

Introduction

Microinjection of cloned DNA into fertilized eggs provides a powerful means for analysing the mechanisms that underly tissue-specific gene expression. Since the DNA is introduced into a totipotent cell, it is inherited by daughter blastomeres that enter many different pathways of differentation. No equivalent test of cell type-specific expression is available using cultured somatic cell lines. Amphibian eggs are especially favourable for such studies since development is rapid and individual tissues are readily isolated from the earliest embryos. It is therefore possible to investigate the regulation of genes activated as the first restrictions on cell fate are established within the embryo.

This procedure was used to study the expression of a muscle-specific gene that is normally activated after the onset of gastrulation. The cardiac actin gene of *Xenopus laevis* encodes one of the earliest products of terminal cell differentiation detected in amphibian embryos (Sturgess *et al.*, 1980; Mohun *et al.*, 1984). Cardiac actin protein is synthesized exclusively in embryonic muscle (myotome) which is derived from the mesodermal germ layer. Previous studies of vertebrate sarcomeric actin genes have indicated that DNA flanking the 5′ end of these genes is required for their expression in myogenic cultures (Melloul *et al.*, 1984; Nudel *et al.*, 1985; Grichnik *et al.*, 1986; Minty and Kedes, 1986). This study tested which region of the *Xenopus* cardiac actin gene is required for activation and tissue-specific expression at the beginning of embryonic development. It was found that microinjected, cardiac actin gene constructs are responsive to tissue interactions that can induce embryonic muscle differentiation from presumptive ectodermal cells in culture. Tissue-specific gene activation and expression in response to inductive stimuli both require a similar region of DNA flanking the 5′ end of the cardiac gene.

Results

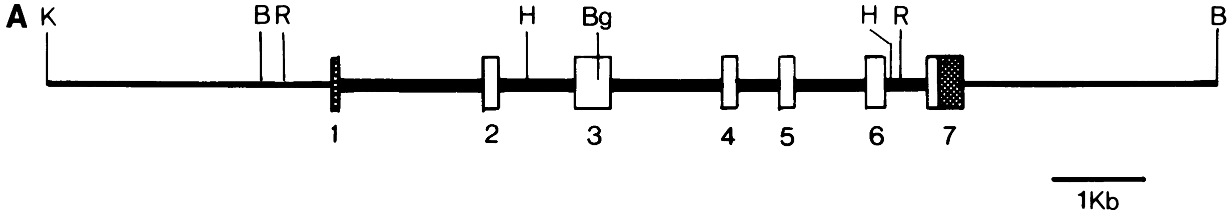
Structure of the Xenopus laevis cardiac actin gene

The cardiac actin gene of *X. laevis* was isolated from a lambda EMBL4 genomic library (Krieg and Melton, 1985) and its structure established by Southern blotting and DNA sequencing. Like all other vertebrate striated muscle actin genes that have been reported, the *Xenopus* cardiac actin gene comprises seven exons (Figure 1A) only six of which encode the mature polypeptide (Buckingham and Minty, 1983). The first exon codes for most of the 5′ untranslated portion of the mRNA. The entire nucleotide sequence of this gene and the flanking DNA is presented in Figure 1B. The positions of each intron with respect to the mature mRNA sequence are identical to those of other vertebrate cardiac and skeletal actin genes (5′ untranslated region, codons 41/42, 150, 204, 267, 327/328). This gene encodes a transcript identical to that of the *Xenopus* cardiac actin mRNA previously reported (Mohun *et al.*, 1984; Stutz and Spohr, 1986).

The transcription initiation ('cap') site was identified by measuring the size of the mRNA 5′ untranslated region using primer extension. This method estimates the transcript leader to be 68 nucleotides long (data not shown) and by comparing the gene sequence with that of the cDNA we predict the first exon to be 51 base pairs long. This is confirmed by RNA mapping (see Figure 6). Upstream from the cap site there is a putative TATA box at −17 (TATAAAT) and a CAAT box sequence (CCAAAT) at −85 characteristic of vertebrate striated muscle actin genes (Ordahl and Cooper, 1983; Minty and Kedes, 1986). A similar but not identical sequence has recently been presented for the first exon and immediate flanking DNA for another *X. laevis* cardiac actin gene (Stutz and Spohr, 1986). The presence of two distinct cardiac actin genes in the frog genome is unusual since all other vertebrates appear to possess only a single copy of each actin gene. However, *X. laevis* and closely related species such as *Xenopus borealis* probably possess a duplicated genome (Bisbee *et al.*, 1977) thereby acquiring a second copy of each single copy gene.

In an attempt to define regions of the cardiac gene which may be important for its tissue-specific pattern of expression, the *Xenopus* sequence was compared with those of other vertebrate sarcomeric actin genes. The results are summarized as follows. (i) No homology was found between exons encoding the mRNA 5′ untranslated regions of the various vertebrate cardiac actin genes. In contrast, Gunning *et al.* (1984) have described several regions of homology shared by vertebrate cardiac actin mRNAs within the 3′ untranslated region. (ii) A short sequence, CACCAG, located at position +7 within the first exon of the *Xenopus* gene comprises the cap site of the mouse, human and chicken cardiac actin genes (Minty and Kedes, 1986). (iii) No regions of homology are evident in the introns of vertebrate cardiac ac-

A



B

GGATCCAATTACAGAAAGACTCCTTATCTAGAAAACACCAGGTCCCAAGCATTCTGCAT -720

ACCAGGTCCCATACCAGTACAATGCAATACTATAAAAGGACACCCTTAGATAATGTACAGATGTCACCTTGGATGCTCGTATAGTGTTGCACTGTGCCCAGATTAACACAATTACCCACA -600

CACATCCTGGGAGAAGGGGGAATTCTGCACATTGGCTGTTTCCCAACAGAGTGCAAGCTCTGAGAGCAAATCTTGCCCGTCATTACTGTTTGATATAAGGTATTCCCATGCCTGACGCAC -480

ATAATATTTTTGTTACATTCGTGTTTACAGCACTGAAGGCAGGCAGCTATTTGTGCAGTCAGTACTTTATGTATTTCTAAACCCAGTGCTGGAGAGGCCAATAGGCAGGATGCCCCCCCC -360

GCGTTAGTCCCCCTGCACAATTGTGCTGCACCTGTCTACTCCATTTGCAGACCCCTGTGTCTGTGCAAACTATTTCTTTCATTGTGCTGTTTTTTTTGTCACCCAGCATTACAGACATGC -240

TTTTTTGGGAATCCCTATTTGGCCATCCCTAGTAGTGCTCCCCCCACCTCCATACTTTCCATACATGGGCTAAGGGGTCCAAAGACCCTGCCCTCCCCCCTCACCTACTCCATTAATGGC -120

TTCTTTGCTTTTCAATGGCCCAGAAGCTACCCAAATAAGGGCAGGCTGCCTGCCTTTCCCTGCGGAGCTCCCACTGACTCCTCAACTCCAGGCAGCGTATAAATTGACAGCTCAGTCCCC

AGGCAGCACCCAGCCAATTCAATCACAGCACCAGCTCTCCTGTGCTAACCAGgtGAGTACTCTGTGGGACATACAGGTTGGTTCTTTGCTGCTGGACGGATGTCTGCCATTTCACTTTTG 120

GGACAAACCAAGATTCCACTGGTGTTTCAGCCATCATTCCTTTAGGATTACCTTTGATTTCCAAATAGTTCCTGGCTCCTATGGCACAGAAAGAGCTATTTAGGATTTCCTCTCTAAACA 240

GTAGCAACTAGTCTCAGAAATGGCTTCAACCATTGGCAGCCTAACTGCTAGTGGATTAAAGGACCCTACAGTTGGTTTCTTGTTTTGCAGAGGAATCTACAATTTCCACTTTAATATGAG 360

GGGGCAGTGTGACCAGTGCACTGTATATAGGGGAGCAGTATATTCCATGTTTATGTACCTGTGTAAGGAGCAACAGCCCCTGGGTGCTACATGGCTCCTATGTATAGACATGTCATGGTC 480

ACAAGTACCTACCTGTATGCTACTAAGGCGCACACACAATTCCTAAAACTGTATTAGAACTTTGAGGCCTACTAAAAATGTGGGTCCCAGTGATACATGGCGTTACTATACAGTAGTACT 600

CACAGGCACGTGCTGGTTATATGATCAGTAATGTGGAACTTGTTGAAGCCCATAGTACACTCCCAATATGTGTTCTGCAGGCTCCTCTGGCAGCAATTTACTTAATGGCAACTCCAACAA 720

ATCAACCCTTACCCCTGTCATTCCTCCAACAGCTGAGGGGGGGGGATAGAGTATTCATCTTTTTACACTGCAGTCATACGAAACTTGTTTTTGCCAGCAGGGTAAGGTTTGTGTGTAGGA 840

GGCAGCTGCTTCTTGACATGTGAGACAGCAGGGTGGTGGCAAATGTTTGGGGTGGGGGGGCATAAACTTTAACTCTTTCAGGTTTAAAGAATAAGTTACACTACTATCACCCTTTAATGC 960

ATTTTTTATTCTTGATTCTAGAACTACTGTTTATACTACATCATATCATACTCTTAGTGCTTGAAAGTTCTTCACCCAAAGTGATAGGGTATTTTGCAGACCCCTTTTCAGTAATACATT 1080

AAATGATTCAATGTTAAATAACATTAAAACAATTCTTTTATTTATCGCCAACATAGCAAGATTGTTAAAGTCATTGTTTTTCCAATTTAATCTAGCTCAGGGATTTTTTCCCTCTGCAAG 1200

TTGGACTTCAATTCCTAGCAGCCCAGGATAGTGGGGGTTGCTGGGAATTGTAGTTCAGCATATTAGTGTAGCAGACATTCTCTGGTGTTGCTGCATATTGGAATTACAGAGATGGGTAGG 1320

TAGACTGGGTTGGTCACTTTTTATAATTCAGCATCTAGACAGCAGCAACAAGCATTACCCTGGAATGGTTCATAATATGCTCAGCTATTATAGTTTCCAAAGTAATAATGTAATATCCTT 1440

GGTCAAAACTACAAAGCCACGTTGCTTTATTTAGACACTCGTTTCTGGAAAGCAAAGGTTAAAAACCAAAGATGTGTGCTCAAAAATCCCTCTCTTCTCTTGCTCACagAACATAACCTG 1560

M C D D E E T T A L V C D N G S G L V K A G F A G D D A P R A V F P S I V G

CCACTATGTGTGACGACGAGGAGACTACCGCCTTGGTGTGCGACAATGGCTCCGGGCTGGTGAAGGCTGGGTTCGCTGGGGATGATGCCCCACGTGCTGTGTTCCCCTCCATCGTTGGTC 1680

R P R H Q

GCCCCCGTCATCAGgtATCTATCCGAAGCACCTTTAGCTCTTTTCTAGACAGGAGGATATAGCTGTACAAGCTTGCAATTTAATGAGCAATACTAGGAAACCTTAAGTAGTTCAGTAGTT 1800

TAGGTGTCTTATCCTAAAATTTGATAAAAATGTTATTCCCCAAATATTCTAACTGCCTATTCCTAGGAATATACTATCATCCACACTCTAGCTCTAGTAAGGCTCTACAGTGCATACTGG 1920

GAGTTCAGAGACAGCTGGAGAGTCGAGTTTGCCAGTCCCTGCCCTTCACCTCTCTATTTGATGCTTTGTTTCCCCTATCAGCAGCTATAAATAGACAGAGTGCATAGCCACTTGTGCAGA 2040

GTCCAGTGATAAGAGCTTGTTAATTATAGAAGAGGCAGACTGTGCCATGTGAATTAGGCCTTATTATTTTGTCAGCATTGGCAGCCAGTGAATGTTTCTGCCATAGCTGACATTTTAGTA 2160

TTATTAATGTCCAGGCTGTTGGGAGTGAAAGCTGAACACATTACAATTACAGCTTTGCCAACTGCCCATTAGTATCACTTTCTGGTGGGTCTTGCCAGTTCCCCCTCTGGGAATGCGATG 2280

GGTCCATTTGCACCTGGGTCAATGGGAGGGAGAATTACTGTATAAAGGGCCAACTAATATTGATGCCAAGTAAGTCATGACCTTGTGTCACCTTCAGTGTCAGGGCAGGAAGACACCATG 2400

G V M

CACTCTGCTTAGATTTATTCAATGGTTTTAAGCAGGAGAGGAATGCTTTTACTTTATGGTTTTGTCTTTGACCAGCTTCTCACCTCGCTTTACTCCTCTCCTTTCTTTagGGTGTCATGG 2520

V G M G Q K D S Y V G D E A Q S K R G I L T L K Y P I E H G I I T N W D D M E K

TGGGTATGGGTCAAAAAGATTCCTACGTAGGTGATGAAGCTCAGAGCAAGAGAGGTATCCTGACCCTGAAGTACCCCATTGAGCACGGTATCATCACTAACTGGGATGATATGGAGAAGA 2640

I W H H T F Y N E L R V A P E E H P T L L T E A P L N P K A N R E K M T Q I M F

TCTGGCACCACACCTTCTACAATGAGCTGCGTGTGGCCCCTGAAGAGCACCCCACCCTGCTCACTGAGGCCCCACTGAACCCCAAGGCCAACCGTGAAAAGATGACACAGATCATGTTTG 2760

E T F N V P A M Y V A I Q A V L S L Y A S G R T T

AGACCTTCAACGTCCCAGCCATGTATGTTGCCATCCAGGCTGTGCTGTCCCTGTACGCTTCTGGTCGTACCACAGgtGAGCATAGACACTTCCATTCAAACCTTAATGATTGGGTTGCTC 2880

TTACCACTGGGAAGGTTCAGGTCTTTCCATCCCTGCTTTGAACTTCATCCTCTGGTTGCTGGTTTAATAAGAAGTGACATAGCTCAAAACTTTCTACATTCTCCTATAGATGGAACTCTT 3000

GTATGAGAAAAGAGGGCAAAATGTATTCTGCTGCTTTGGAACAAACCCTCACTATTACCAGTAATTCCAAAAGCTGGAGATAAATAGAGACAAAATATATATTCCTAGGTTGAATTTTTA 3120

TCTAAAACCTTCCTTGAAAGCAAACCAGCTAGTGGATCTTCCTTTGCACTTATATTTATCAGGCACCCCAGAGATACTACAAAGAACCATACAACCTACTGTGCTCCTGTGCAAGTGCTG 3240

GTATCTACATTGTGTATCTGAAGTCAATGGTTAACATTTATAGAAACCTTGCCCTTCATCCCCATCCAGAAGAGAATGGCTATAAATTAGCTATGAACTCTCTTCCCAGTAATACATCAA 3360

AGTCACATTAATTCATCCCAAGTTCAAGCAGAATGCAGTGGGAGCTTTAATCTTTTTACACTGGTAACATCTCATCACTGAATCCAGCCTATTTAGTTGTAAGGCAATGCTGGGTTTGTA 3480

TGTACCAGAACAGCACATCTCCAATGGAGGGTGGGGGGAGCGAGGGAGTAAAGGGGCATCAGCATCTGCCTGCTATATAAATAGAGGGGGAAAATGTGACATTGTGTGAGCTCTGCATAT 3600

TTTTTTTCAACGTTAAATCCCTTAAAGCTATTGGTATGATAAGAAAGCGGAAAGGCACCATGAGCGTAATAAGCCTTCTCATACAGTCTGCAGGCTGAGATTGCTTGCCAGCAGATACGT 3720

TGGATTATTTCTGCTCATTGCAGTGATGTCTCTTGCCCTTCAGGACAGAGGCTTTCTGTGTTTACATGCAATTTATTTGAGTTTAGATCCACAAAACACAGCTTGGCACATCACTCTCTG 3840

TCTTCAGCCTGGATAGTGTCCAACTTGCTATCTGACTCTCTGGGAAGCTCAGTTCCCAGCACTATTTAATAGAGTTGAAGTTGGATCTTTCTTCTGTAGAATAACCTTTGCTAATTAGGC 3960

G I V L D S G D G V T H N V P I Y E G Y A L P H A I Q R L D

CTTAACATTCATCTATTCTTCTTTATCCagGTATCGTTCTTGACTCTGGTGATGGTGTCACCCACAATGTCCCCATCTATGAAGGTTATGCTCTGCCCCATGCCATCCAGCGTCTGGACC 4080

L A G R D L T D Y L M K I L T E R G Y S F V T T

```
TGGCTGGTAGAGACCTCACAGATTACCTCATGAAGATCCTGACTGAACGTGGCTACTCCTTTGTGACAACAGgtAAGTTCTATCATGCTAAATCCATAGAGGGCCTACACAAACTTAGAT  4200
AGCATTTGCCGAAGAATACAGAATCTATTATATACACTTGGAAGAAAATTATGTCATTATAAGACAACAAGAAACAGTGACAGTCACAGACTGATACATCAGCTGGGCTATGCACTAATT  4320
ATTGAACCTTGTGATATGTAGCAATTATGTCTTATAAAAAAGTCATAGGACCCCCGGGCACATACCGAAAAATAACTCCCCCATAGATGAGAATCTCACGTAGAATCATTGAATGAGGAC  4440
ATTTACTGTGACCACAAAGCAGAACATCTTACTAATAAGAGAGAAAATAGCCACAATACTGAAAATAATGAACTTGTGATTTTTTTCAATGTTTCTGTAGAATAACTCTTCAGAGTTTAA  4560
                                    A  E  R  E  I  V  R  D  I  K  E  K  L  C  Y  V  A  L  D  F  E  N  E  M  A  T  A  A  S  S
TCTCATTATGCTTTGTTTTTGCCCCATACagCTGAAAGGGAAATTGTCCGTGACATCAAGGAAAAGCTGTGCTATGTGGCTTTGGACTTTGAGAATGAAATGGCCACCGCTGCCTCATCC  4680
 S  S  L  E  K  S  Y  E  L  P  D  G  Q  V  I  T  I  G  N  E  R  F  R  C  P  E  T  L  F  Q  P  S  F  I
TCCTCCCTGGAGAAGAGCTATGAGCTTCCCGACGGTCAGGTCATCACCATTGGAAATGAACGTTTCCGTTGCCCTGAGACCCTCTTCCAGCCTTCCTTCATTGgtGAGCAGCCATACCTA  4800
TAGCCCTGAGACCTGACATATACCAACACAGATACACAGACTCCTGTTTAGTGATGGGTGAATATATTCGCGAACCGTCGCGAAAATTTGTTTTTTGGATGCCGGTGCAAAAAACGGATG  4920
CCGGCGTCCAAGAAATGGACTTCGGTGTCAACAAACAGATGCCGTTTGTGAATTTTTCAGCATTTAGAGAATTTCGCAAATTTTTCGGCGAACGGAAATGCCCCAAATGCGCCCATCACT  5040
ATTCCTGTTCAGTTTTCCCTATACCACTTCACATACTAGAAATCCATATCATAGGGCAATATGCGTTTTATATCACAGTGTGCCTACATAAGCTTCTCATTAACCTTTTACTTTTTTACT  5160
AACCAGAGCGAAACCCCCCTAGGTAGACAAAATATATTTAAGATATACTGTAATATTAGAGACGTGCCTGGTTAAGGCCAGTCTCATTCTAGATCCTGTAAAAAATATTCACACCATCTTC  5280
CCTTGCCAACCAACACATGCTGCCGTTGCCTTAATAAAGCCCTATTTATCATTTGTTAGACCCCAACACCTGTTTTCAGTGACGTGCCCACACAGCAGTTGCTTTCCATTTGGTGATGTT  5400
                                                                                                            G  M  E  S  A  G  I
GTTTCAGCCTCACACTCCCATCTACACCTGTTCCCCCCACATTCCCCCCACACAGTTATGCCTCTCTAATGCATCCTGTTTTACCTTCTCCCACTACAagGTATGGAATCTGCCGGTATC  5520
 H  E  T  T  Y  N  S  I  M  K  C  D  I  D  I  R  K  D  L  Y  A  N  N  V  L  S  G  G  T  T  M  Y  P  G  I  A  D  R  M  Q
CATGAAACTACCTACAACAGCATCATGAAATGCGACATTGATATCCGCAAGGACCTGTACGCCAACAACGTCCTCTCTGGTGGTACCACCATGTACCCTGGTATTGCTGACAGAATGCAG  5640
 K  E  I  T  A  L  A  P  S  T  M  K  I  K
AAGGAGATCACTGCCCTGGCCCCCAGCACCATGAAGATCAAGgtGAGCTGAGCAAATGACAGACCTACAATAGACATTTCTTACACTGTATATGCTTCAAATTTGATATAAGAGAATTCA  5760
GAAACATGTGCCCACTGCCCAATCACATTCTTCTTATGTTTGTATACTATACATTTCCCTGGGCAGATATTCCTCTCCTCTTCTATAGAGTAACACTTCTCATGAACACAAATATTCACT  5880
TGTTACTCTTCCTCTGATCACAATGATTATTATATCAGGAATGGAAACATAGGAGCCCCAAACAGGGGAGAGTGGCACCCATCAGGACATATAATGAGTAAATCTGTAGTACATAATGAA  6000
CAAGATGCGTGCAATACAGAATAAGTTGAAGTTGTGCTGTTAGATACAGGCTTGTGGTAGAGGGTTGCATCTCAACAGGTCTTTGCTCATGAATTTATTTTAATCAATTATTTCCTCTag  6120
 I  I  A  P  P  E  R  K  Y  S  V  W  I  G  G  S  I  L  A  S  L  S  T  F  Q  Q  M  W  I  S  K  Q  E  Y  D  E  A  G  P  S
ATCATTGCCCCACCTGAGCGTAAATACTCCGTCTGGATCGGTGGCTCCATCTTGGCCTCTCTGTCCACCTTCCAACAGATGTGGATCAGCAAACAGGAATACGATGAGGCAGGACCCTCC  6240
 I  V  H  R  K  C  F  *
ATTGTGCACAGGAAGTGCTTCTAAACATTACACACACTCCTCCAAGCAACATATACAGAGTACCTCAAAGTCAAAGAAAACAAGACCTGGACAGTCACTGTACAGTTTGTTTACACGCCA  6360
TCATATTTATTACTTTCTAAATAAAACACATGCGGTTAAGTTCATCAGAGTGTGACTGTCTTCTTTTGGATGTTATGCTTTGCCAATAAAAGGAAACATCGGCTCAACCAT

                                                                                                                GGTTTCGGC  6480
AGTTGTGTGTGACAGTTTATATTACAAATAGGAGTTTTAATGAGGTTTGGGGGCTGAAATAATAAAGTAATAAAGTAATAAAGTAAAGGTGCTTCAAACGATCAATAAGTATAGATGTTT  6600
AGGATCAAACCAAGGGGCTTACTGTAAAATACATTCAACATCACATGAGAACTAAGCCAGACAATATGTCATCAGCACTTGCTGAATAAAAGGCAACTTTAGCATTATCTGGCCTCTGAG  6720
AAAGAAATACAGGTATGGTATCGATTATCCGGAAACCCAGAGAGCTCTGAATTACAGAAAGCCTGTCTATCATAGTCTCCACTTTATCCAAATAATCCCATTTTTTAAAATGTCCTTTTT  6840
CTCTGCAGTAATAAAATAATACCTTGTACTTGATCCAAATTAGGATAAATGAATCCTTATTGGAGGGAAAACCAACCTATTGAGTTTACTCAATGTTTACATGATTTCTATGAAGATCA  6960
AAAGAGAGATTCGATTTCTGAAAAAAATCCCAAGTCCCGAGCATTCTGGATAAAAGTTCCCATACCTGTATATTTGCCTAAACATATTAGATTTGGAGGCTATGATGACAAATATGTGAGT  7080
GCGAGCATTTCCCAGAACACTAATATTCCTTTGAAGAACCTCAACACCATGAATCCAAAATAAATAATGAGAACACTCATCTTGTTAACAACAGCCATGCAAAATAGAACAGGTGGTTAA  7200
ATGCACATACACATACTAAAGCTAAGGAGTGCCAAACAAGAAAGGGTTCGACAAACAAAGGGAACAATTAACCTTCTAGGATCAAATACCCTACAAACATGAAGTTGTCAGCTCAAACAG  7320
TAAAAGAGCAATAATCAACTACATTAATTTGCACCAATTGAGTCACACTTAGGGACAGATTTACATATGGTGGAATATCAAGGGTTAATTAACCCTCGATATTCGACTTCCGAAGTCGAA  7440
GGATTTACCACAATTCGTTCAATCGAACGGTTAAATCCATTTTAATCCATCGATCGAACGATTTTTCTTCGACAAAAAAAGATTGCAAAGCCTATGGGGACCTTCCCCATAGGCTAAGAT  7560
TGACTTCGGTAGGTTTTAGGTGGCGAACTAGGGGGTCGAAGTTTTTTTAAAGAGACAGTACTTAGACTATGGTCAGTACTTAGACGAATGGTCGAATAGTCGAACGATTTTTCGTTCGAA  7680
TCGTTCGATTCGAAGTCGAAGTCGTAGTGGAAGGTTGAAATAGCCCATTCGATGGTCGAAGTAGCCAAAAAAAATTCGAAGTTTTTTTTATTCTATTCCTTCACTCGAGCTAAGTAAATG  7800
GGCCCCTCTCTGTATAATGATGCAGTTCATATAGCACCAATATATTCTGCAGAATTGTACAGAAATGATACATTATTCACACCAATCCCTGCCCCAGTGGAGTTTATAATCTACAGTCTC  7920
TATAAATATTCATACACACCTTATAGTCAATATTATCTGGAGACATCTAACCTGTATGCTCTAGGAGTGTGGGAGCAAAACCCCACACATCTAAACAGCACAGAGGCACCAATTTTAAATG  8040
TTTTTATTCCTATACTACATCATATTCCTAGTTTAAACTCTACGTGTGTATTGAAACTTCGCTTATTAACCAGGTGAGGCTGTGCACATA
```

Fig. 1 (A). Structure of a *X. laevis* cardiac actin gene. Boxes indicate exons (untranslated regions cross-hatched), numbered 1 – 7. The thick line represents introns. Restriction sites: **K,** *Kpn*I; **B,** *Bam*HI; **R,** *Eco*RI; **H,** *Hin*dIII; **Bg,** *Bgl*II. (**B**). Nucleotide sequence of the *X. laevis* cardiac actin gene shown in (A). The complete sequence of the gene is shown (along with flanking DNA) with nucleotides numbered from +1 at the start of transcription. Intron boundaries are shown in lower case. Putative TATA and CAAT boxes are underlined, as are the two hexanucleotide, polyadenylation signals in the final exon. The deduced cardiac actin amino acid sequence is shown in single letter code. The termination codon is indicated (*).

tin genes. (iv) Neither is their any similarity between any portion of the *Xenopus* cardiac gene and chick (Ordahl and Cooper, 1983; Nudel *et al.*, 1985) or rat (Zakut *et al.*, 1982) skeletal actin genes other than their common CAAT box. (v) Within the upstream sequence flanking the vertebrate cardiac actin genes Minty and Kedes (1986) have recently identified a common, repeated sequence motif comprising CC(A/T)nGG an example of which is the striated muscle actin gene CAAT box (CCAAATAAGG). Such motifs are also present in the *Xenopus* cardiac actin gene promoter region (Figure 2).

To test the role of such conserved sequences, derivatives of the cloned cardiac gene were introduced into fertilised eggs and their subsequent expression in early embryos was studied.

A chimaeric cardiac actin-bacterial chloramphenicol acetyl transferase (CAT) gene is correctly regulated in early embryos

Linear, plasmid DNA persists after microinjection into fertilized eggs and is retained in an episomal, concatenated form in the nuclei of the embryonic blastomeres at least for the first day of development (Rusconi and Schaffner, 1981; Bendig and Williams, 1983; Etkin *et al.*, 1984). During this period, extensive cell differentiation occurs to provide the major embryonic organs and tissue types of the larval tadpole. It has been shown

```
                 **      **
A   -230  ATTCCCTATTTGGCCATCCC                 Xenopus
    -240  GCTCCCTATTTGGCCATCCC                 Human
    -241  GCTCCCTATATGGCCATTGC                 Mouse
    -197  GCTCCCTATTTGGCCATGTG                 Chicken

                            **        **
    -197  CCCACCTCCATACTTTCCATACATGGGCT        Xenopus
    -212  CCC.CCTCC...CCTTCCTTACATGGTCT        Human
    -213  CCC.CCTCC...CCTTCCTTACATGGTCT        Mouse
    -168  CGCTGCTCCTCACCTGCCTTAGATGGCCG        Chicken

              **       **
    -132  CTCCATTAATGGCT                       Xenopus
    -151  CTCCATGAATGGCC                       Human
    -152  CTCCAAGAATGGCC                       Mouse
    -121  GGCCATTCATGGCC                       Chicken


B   -323  TACTCCATTTGCAGACCCCTG           Xenopus cardiac actin
    -71   TTAATAATTTGCATACCCTCA           Mouse Ig V_L
    -63   CACATGATTTGCATACTCATG           Mouse Ig V_H
          TGGGTAATTTGCATTTCTAAA           Mouse Ig intron (V-C)
    -256  CACCCTATTTGCATAGCCCGC           Xenopus U2
    -261  CAGACTATTTGCAA                  Xenopus U1
```

Fig. 2. (**A**). Evolutionarily conserved sequences in the upstream regions of vertebrate cardiac actin genes. In addition to the highly conserved, sarcomeric actin CAAAT sequence (Ordahl and Cooper, 1983), there are several highly conserved sequences, as shown. The location of each with respect to the transcription initiation site is indicated. Each conserved sequence contains a CCarGG motif (marked by asterisks). The human, chicken and mouse gene sequences are from Minty and Kedes (1986). (**B**). Presence of sequence homology in the *Xenopus* cardiac actin gene with an identified transcriptional control element of other vertebrate genes. Sequences from the upstream regions of the mouse immunoglobulin heavy (Ig V_H) and K-light chain (Ig V_L) are aligned with respect to a common octamer (ATTTGCAT) transcription factor binding site (Singh *et al.*, 1986). A similar sequence is found in the immunoglobulin enhancer element located in the intron between variable and constant regions of heavy and light chain genes (Ig intron V-C), and in the enhancer-like sequence upstream of the *Xenopus* U1 U2 SnRNA genes (Mattaj *et al.*, 1985). The location of each sequence is indicated.

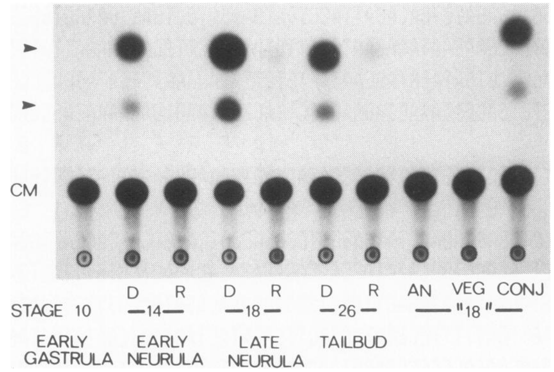


Fig. 3. CAT activity assays on embryos injected with plasmid DNA containing the cardiac actin gene promoter (3 kb) fused to the bacterial CAT transcription unit. Embryos were injected with DNA at the two-cell stage and cultured whole until the stages indicated. A dorsal, axis-containing fragment (**D**) was dissected from the rest (**R**) of the embryo and each part analysed separately. (This dissection is not possible prior to gastrulation.) Other embryos were dissected into animal (**A**), vegetal (**V**) and equatorial pieces (as described by Gurdon *et al.*, 1985) at stage 8, and the fragments cultured until stage 18. An. and veg. pieces were also placed in conjugation (**C**) at stage 8 and cultured until stage 18. **CM**, non-acetylated chloramphenicol. Arrows indicate the mono- and di-acetylated forms of chloramphenicol.

previously that the cardiac actin gene is activated along with its skeletal actin counterpart in the presumptive muscle tissue of the embryonic mesoderm after only 20 h of development (Mohun *et al.*, 1984). It is possible therefore to examine whether microinjected copies of the cloned cardiac gene are appropriately regulated in a temporal and tissue-specific manner in the developing embryo. Figure 3 shows the results of such an experiment.

To distinguish between expression of the endogenous cardiac actin gene and that of the microinjected copies, a chimaeric gene was constructed comprising 3 kb of *Xenopus* DNA including the cardiac actin gene promoter and cap site attached to the entire bacterial CAT gene (Gorman *et al.*, 1982). The fusion point lies at position +24 in the cardiac actin gene, within the 5′ untranslated region of the mRNA. When this chimaeric gene is introduced into fertilized eggs no CAT enzyme activity is detected for the first few hours after injection. The cloned gene remains inactive until gastrulation (stage 10), after which enzyme activity can be detected in the dorsal, axial portion of the embryo. This region contains the presumptive somite tissue that forms the embryonic muscle in addition to other axial structures such as notochord and neural tissue. Little or no enzyme activity can be detected in the remaining tissues of the embryo though these contain more cells than the dorsal fragment. As development proceeds through neurulation (stages 14–18) and the first morphologically distinct somites are established, the level of CAT activity increases approximately 7-fold in the dorsal region of the embryo. From this stage onwards, the expression of the chimaeric gene varies considerably between individual batches of embryos; in some the level of CAT activity continues to increase in tailbud embryos (e.g. Figure 3); in others, the level declines several fold. Earlier studies have indicated that the survival of microinjected DNA is highly variable in neurula and tailbud stages (Newport and Kirschner, 1982; Gurdon and Smith, 1984) and that the amount of transcription from injected genes corresponds to the number of copies which persist in the embryo (Rusconi and Schaffner, 1981).

Although the chimaeric gene is subject to correct temporal regulation, the region-specific location of CAT enzyme activity could be the result of differential stability of the bacterial enzyme in the various embryonic tissues. It could also reflect a mosaic distribution or survival of the microinjected genes amongst the early embryo blastomeres. These possibilities were tested in two ways. Firstly the expression of the chimaeric gene in isolated, cultured regions of the early embryo both before and after their conjugation was examined. The equatorial portion of a blastula will form axial structures containing muscle tissue when cultured separately (reviewed in Slack, 1983). Consistent with this, it was found that copies of the microinjected fusion gene are activated in the equatorial but not animal or vegetal regions after several hours of culture (Figure 3). This is in agreement with the transcriptional activation of the endogenous cardiac actin gene in similar regions after isolation and culture (Gurdon *et al.*, 1985). Secondly, it is well established that the animal pole region of the blastula embryo, which gives rise to ectodermal tissues during normal development, may be induced to form mesodermal derivatives including muscle tissue by contact with vegetal cells of the blastula (Nakamura *et al.*, 1970; Sudarwati and Nieuwkoop, 1971). Conjugates formed from these two embryo

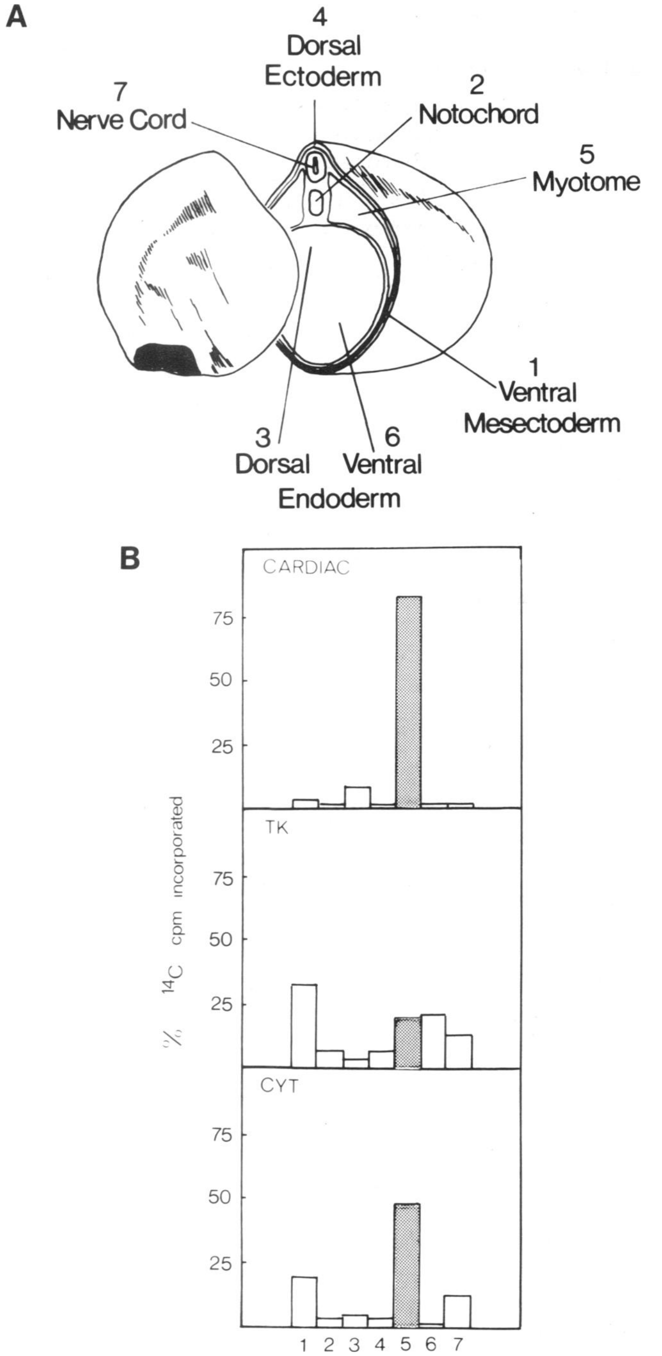


Fig. 4 (**A**). Diagram of a stage 18 embryo, seen in traverse section, with the head fragment on the left. The regions indicated (1 – 7) were isolated by dissection and immediately frozen. (**B**). Histograms summarizing the results of assaying CAT activity in the dissected regions of embryos indicated in **A**. The enzyme activity of each fragment is expressed as a percentage of the total for all seven regions of the embryo. Embryos were injected with DNA containing the bacterial CAT transcription unit fused to upstream sequences encompassing the promoters of the *Xenopus* cardiac actin, Herpes simplex TK and *Xenopus* cytoskeletal (type 5) actin genes.

fragments show levels of enzyme activity comparable with undissected embryos after culture until control embryos have reached the neurula stage (Figure 3). This activity is located exclusively within the animal fragment of the conjugate (data not shown). The microinjected fusion gene therefore responds to inductive tissue interactions in an identical manner to the endogenous sarcomeric actin genes (Gurdon *et al.*, 1985). Since none of the tissue in the conjugates would normally contribute to muscle tissue

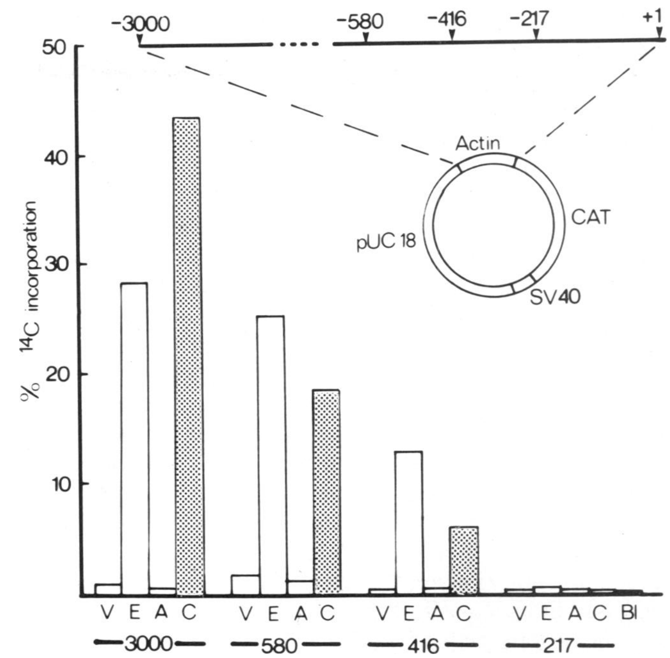


Fig. 5. Summary of CAT assays on regions of embryos injected with 5′ deletion mutants of the cardiac actin-CAT chimaeric gene (shown). Enzyme activity is expressed as the percentage conversion of chloramphenicol to its mono- and di-acetylated forms, all assays being performed under identical conditions. Injected embryos were dissected at stage 8 into vegetal (**V**), equatorial (**E**) and animal (**A**) fragments. Conjugates (**C**) were also made from animal and vegetal fragments. Embryo pieces were cultured until stage 18 and then assayed. The extent of cardiac actin gene upstream sequence remaining in each deletion mutant is indicated beneath each histogram.

it may be concluded that the presumptive ectoderm tissue of the early embryo retains sufficient copies of the chimaeric gene to permit ready detection of the bacterial enzyme after induction. The predominantly dorsal expression of the injected gene in neurula embryos is unlikely therefore to be the result of differential distribution of the injected DNA within the embryo.

Tissue-specific expression of the cardiac actin-CAT chimaeric gene

To examine the precise localization of expression of injected genes amongst the tissues of the early embryo, seven different regions dissected from neurulae were tested for the presence of CAT activity (Figure 4). Embryos injected with the cardiac actin – CAT gene show 80 – 90% of their activity within the myotome tissue. A low level of activity is detected in the ventral mesectoderm and the dorsal endoderm. In both cases it is probable that this results from contamination of these fragments with myotome tissue during microdissection. Previously, a low level of endogenous cardiac actin mRNA was found in the ventral mesectoderm and was considered an artefact resulting from an arbitrary delineation between ventral and dorsal mesoderm (Mohun *et al.*, 1984).

This dissection was repeated on embryos injected with another chimaeric gene, comprising the bacterial CAT gene under the regulation of the Herpes simplex thymidine kinase (TK) promoter. In this construct, the position of fusion lies at +52 with respect to the TK cap site, upstream of the TK protein-coding sequence. The chimaeric gene is flanked by approximately 700 nucleotides of TK sequence including its well-characterized promoter (McKnight and Kinsbury, 1982). Figure 4 shows that the TK – CAT gene is expressed to varying degrees throughout all the tissues of the neurula embryo. In contrast to embryos injected with the cardiac actin-CAT gene, those receiving the TK – CAT fusion show no elevation in enzyme activity within the myotome tissue.

Figure 4 also shows the tissue distribution of CAT activity in embryos injected with the CAT gene under the regulation of a 523 nucleotide fragment encompassing the *Xenopus* cytoskeletal actin gene promoter (T.M. unpublished results). Again the point of fusion lies within the first exon of the actin gene which encodes the 5′ untranslated region of the cytoskeletal actin mRNA. This mRNA, which encodes a γ-like cytoskeletal actin protein is inherited by all tissues of the early embryo from the maternal stores of the oocyte. *De novo* synthesis of the message commences at gastrulation to varying extents in all tissues of the embryo (Mohun *et al.*, 1984; T.M. unpublished results). Similarly, the cytoskeletal actin – CAT chimaeric gene is expressed in a correct temporal manner (data not shown) and varying levels of enzyme activity are detected in different regions of the embryo. Interestingly, the myotome tissue contains considerably more CAT activity than any other and this provides the major difference between the profiles of TK – CAT and cytoskeletal actin – CAT gene expression (Figure 4). Since both these fusion genes are expressed in most parts of the neurula embryo, it is concluded that at least some cells of each region of the developing embryo receive and retain copies of the microinjected gene. Furthermore, the CAT enzyme product common to all three chimaeric genes is stable throughout the embryo. The cardiac – CAT fusion gene may therefore be considered to be correctly regulated in both a temporal and tissue-specific manner in the early *Xenopus* embryo.

Upstream sequences required for correct expression of the cardiac actin-CAT chimaeric gene

The cardiac actin – CAT fusion gene contains only 23 nucleotides of the cardiac actin mRNA and yet the gene is regulated in parallel with the endogenous cardiac actin gene. This suggests that no portion of the actin gene downstream from position +24 is essential for temporal- or tissue-specific activation during early embryo development. It was therefore attempted to define the extent of sequence upstream from the cap site that is required by the fusion gene for appropriate regulation in early embryos. By injecting 5′ deletion mutants of the chimaeric gene into fertilized eggs, the effect of increasing truncation of the *Xenopus* promoter fragment on expression of CAT activity was tested.

Two types of experiment were used to assess whether the chimaeric genes are expressed in a tissue-specific manner. In the first, injected eggs were allowed to develop until the neurula stage and then dissected into dorsal (axial) and ventral portions. Each fragment was assayed for CAT activity. In the second, blastulae from injected eggs were dissected to obtain isolated animal and vegetal fragments neither of which would express muscle-specific genes when cultured on their own. These were recombined and the resulting conjugates tested for CAT activity after a period of culture (see above). Both approaches gave similar results and an example of the latter is shown in Figure 5.

Two results are striking in these experiments. Firstly, any expression of the fusion gene requires sequence(s) lying between 217 and 416 nucleotides upstream from the start of transcription. This region is required both for the induction of the fusion gene in animal/vegetal conjugates (Figure 5) and for myotome-specific expression in the intact embryo (data not shown). Secondly, the absolute level of CAT activity found in myotomes or in conjugates declines with increasing truncation of the cardiac actin promoter from −3000 to −416. It seems unlikely that this results from the increasing proximity of vector sequence to the fusion gene cap site since the injected gene is linearized at the 5′ junction of pUC 18 vector and *Xenopus* DNA. Furthermore, Wilson *et al.* (1986) have found that the concatenates of injected DNA found in early embryos contain monomers in random orientation. One possible explanation is that the quantitative level of transcription from the cardiac actin gene is modulated by several sequences dispersed through the upstream DNA. These appear distinct from the sequence(s) between −217 and −416 that confer tissue-specific expression on the fusion gene.

The cardiac actin gene is transcribed in oocyte nuclei

Since the cardiac actin gene promoter is capable of conferring tissue-specific expression on the bacterial CAT gene it was expected that the fusion gene would, like the endogenous sarcomeric actin genes, remain quiescent in oocyte nuclei. Indeed, the linearized fusion gene gave little expression in oocytes, however, this could well have resulted from degradation of the microinjected DNA during the 24 h incubation period. As some genes microinjected in supercoiled form are correctly expressed in oocytes (Bienz, 1984) the experiments were repeated using circular plasmid DNA. Surprisingly, the entire series of 5′ deletion mutants was expressed after microinjection into oocyte nuclei whilst the promoterless CAT gene gave rise to no detectable enzyme activity (data not shown). Truncation of the cardiac actin promoter to −56 with the concomitant removal of the cardiac actin CAAT box had no effect on the level of CAT enzyme found in the oocyte cytoplasm. Expression of the fusion gene in oocytes does not therefore require the same upstream sequences necessary for embryonic expression.

No cardiac actin mRNA is detectable in uninjected oocytes or very early embryos (Mohun *et al.*, 1984) and it is thereby possible to examine directly the transcription of an entire cardiac actin gene introduced into oocyte nuclei. RNase protection was used to study both initiation and termination of cardiac actin gene transcription in microinjected oocytes. Using a probe that spans the 51-nucleotide first exon of the gene, it was found that only a few percent of transcripts are correctly initiated and spliced in the oocyte germinal vesicle. The majority constitute readthrough transcripts that appear to be initiated at a few preferred sites in the promoter region (Figure 6). Curiously, as the amount of injected DNA increases, the level of most readthrough transcripts declines and the proportion of correctly initiated increases several fold to a plateau level. Similar results are obtained with the cardiac actin gene cloned in either orientation in both pUC and pBR322 plasmids and is unaffected by truncation of the 5′ flanking sequence to as little as 200 nucleotides. The gene is correctly transcribed 10-fold more efficiently when introduced into oocytes in a lambda phage vector (data not shown).

Cardiac actin transcripts are correctly processed in oocyte nuclei

Finally, the cardiac actin transcripts detected in microinjected oocytes were examined for possession of a correct 3′ terminus. Using a probe that spans the last exon of the cardiac actin gene and includes flanking DNA sequence, a single major protected fragment of approximately 280 nucleotides was found. This maps the 3′ end of the transcripts to a position immediately downstream from the first putative polyadenylation signal (Figure 1B) and is the same for the endogenous cardiac actin mRNA of early embryos (Figure 7). A small proportion of transcripts in RNA from both embryonic muscle and microinjected oocytes possess a longer 3′ untranslated region which terminates at two positions within 20 nucleotides downstream of the second polyadenylation signal. cDNAs corresponding to both sized mRNAs have

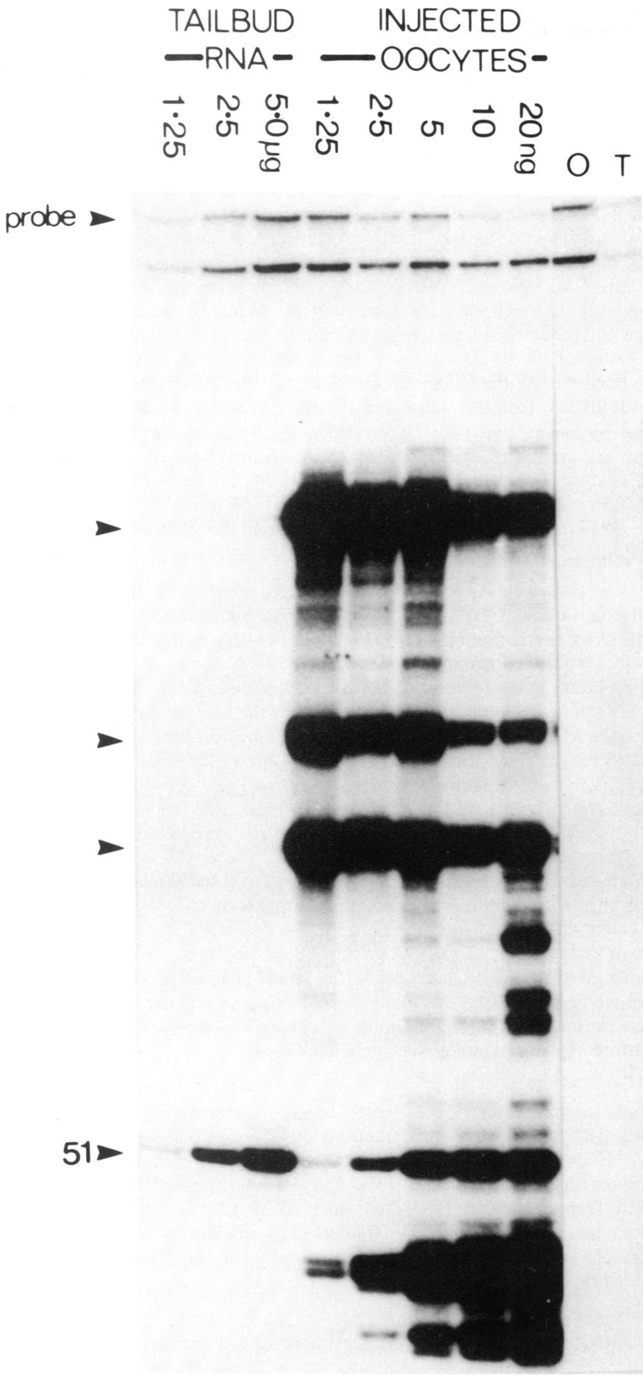


Fig. 6. Transcript initiation analysis of RNA from oocytes injected with a plasmid containing the entire *Xenopus* cardiac actin gene, including 3 kb of upstream sequence. RNase protection of a 51 nucleotide fragment represents correct initiation of cardiac actin gene transcripts, as seen in normal tadpoles (left three lanes). **Lanes 4–8** show the results with oocytes injected with 1–20 ng of DNA; unlabelled arrows indicate major protection products resulting from incorrect initiation of transcription of the plasmid DNA. **O**, uninjected oocytes; **T**, tRNA. 2.5 µg of total RNA was used for each lane.

been isolated from RNA confirming that both polyadenylation signals are used during early development (T.M. unpublished results). From Figure 7, it appears that similar levels of correctly processed transcripts are obtained from oocytes injected with 1.25 and 2.5 ng of DNA despite the several fold difference in the levels of correctly initiated transcripts in the same samples (see Figure 6). This suggests that at least some of the readthrough

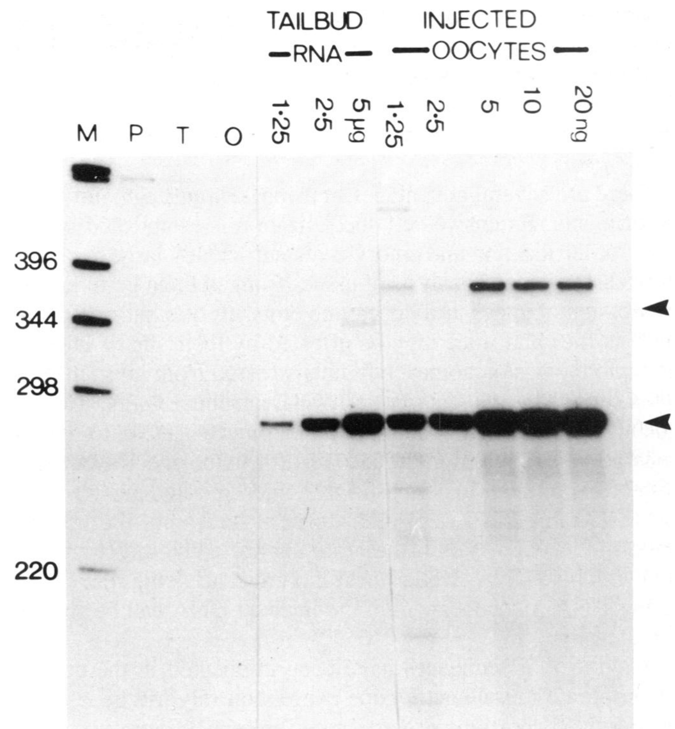


Fig. 7. Analysis of transcript 3′ termini in RNA from oocytes injected with the *Xenopus* cardiac actin gene, as in Figure 6. The lower arrowed band indicates a 280 nucleotide fragment that represents the normal 3′ terminus of the cardiac actin gene transcript. A small proportion of cardiac actin mRNA in tadpoles possesses a longer 3′ untranslated region and gives a doublet of protected fragments (upper arrow). **M**, DNA size markers; **P**, undigested probe; **T**, tRNA; **O**, uninjected oocyte RNA. 2.5 µg of total RNA were used in each assay.

transcripts are processed to obtain a correct 3′ terminus and are sufficiently stable to be detected by RNA mapping.

Discussion

All three classes of eukaryotic genes are transcribed after injection into fertilized *Xenopus* eggs. Those utilizing polymerase III such as 5S and tRNA genes are efficiently transcribed and their products readily detected (Gurdon and Brown, 1977). Polymerase I-transcribed ribosomal RNA genes are also expressed in the early embryo, transcription being activated at the late blastula stage in parallel with their endogenous counterparts (Busby and Reeder, 1983). In contrast, only a relatively low level of expression has been reported for protein-coding genes derived from organisms as diverse as the fruit fly (Etkin *et al.*, 1984; Etkin and Balcells, 1985), sea urchin (Bendig, 1981; Etkin *et al.*, 1984) or rabbit (Rusconi and Schaffner, 1981) after their injection into *Xenopus* eggs. In no case has correct temporal regulation of such genes been obtained. Even with *Xenopus* globin (Bendig and Williams, 1983, 1984) or vitellogenin genes (Andres *et al.*, 1984) developmental expression is incorrect in early embryos. Recently, however, faithful temporal expression of a *Xenopus* gene, GS17, has been reported by Krieg and Melton (1985). This newly characterized gene is normally activated at the mid-blastula stage and transcribed through the first stages of gastrulation. It has been suggested that since injected DNA is frequently lost from embryos after the end of gastrulation, correct temporal regulation is only likely with genes (such as GS17) that are normally activated in very early development. The results reported here are in agreement with this, the cardiac actin–CAT fusion gene showing a temporal and spatial pattern of expression in embryos identical to that of the endogenous sarcomeric actin genes. Similar results have also been obtained by Wilson *et al.* (1986) with an entire cardiac actin gene from *X. borealis* introduced into eggs of *X. laevis*.

There are several advantages in using *Xenopus* eggs for these experiments. Extensive cell specialization is established within a day of fertilization and embryos are sufficiently large to enable microdissection and analysis of tissues from all three germ layers. Expression of genes introduced into embryos of other organisms such as flies and mice may be affected by their site of integration into the host genome. Individuals reared from injected eggs often show a mosaic distribution of the introduced gene. Consequently, germ line transmission of the injected DNA to the F1 generation is required in order to test for tissue-specific patterns of gene expression. In contrast, DNA microinjected into *Xenopus* eggs is retained in an episomal form at least until the neurula stage and its expression is therefore unaffected by direct interaction with host DNA. Under these circumstances, the role of *cis*-acting DNA sequences within the injected DNA can be studied more easily.

A number of sequences have been implicated in the control of vertebrate muscle actin gene expression. Hybrid genes containing the 5′ portions of vertebrate sarcomeric actin genes are activated during terminal differentiation of myogenic cells. A region of 5′ flanking DNA within a few hundred nucleotides of the cap site is sufficient to confer this pattern of expression on fusion genes containing the first exon of the cardiac and skeletal actin genes (Nudel *et al.*, 1985; Grichnik *et al.*, 1986; Minty and Kedes, 1986). 5′ deletions have been used to identify two distinct regions flanking the human cardiac actin gene that are required for expression in a differentiating myoblast cell line. The precise sequences involved in this control are unknown. In the present study, it has been shown that a similar restricted region flanking the *Xenopus* cardiac gene is required for both temporal and tissue-specific expression in the earliest stages of embryo development. It is noteworthy that this region, spanning from −217 to −416 includes a short sequence (Figure 2) that closely resembles both the *Xenopus* U2 SnRNA enhancer (Mattaj *et al.*, 1985) and a transcriptional control element of mouse immunoglobulin genes (Singh *et al.*, 1986).

Microinjection of genes into *Xenopus* oocytes has provided an alternative to *in vitro* transcription systems in the identification of *trans*-acting factors regulating gene expression (Galli *et al.*, 1983; Ferguson *et al.*, 1984). For this reason the expression of the cardiac gene in oocyte nuclei was examined. Only a small proportion of transcripts are correctly initiated and it may therefore be possible to establish a 'complementation' assay for extracts of myotome or muscle tissue capable of stimulating accurate transcription of the injected gene.

Materials and methods

Isolation and characterization of the cardiac actin gene

Several recombinant phage were isolated from a *X. laevis* genomic library kindly provided by D.Melton using a gene-specific probe derived from the *X. laevis* cardiac actin cDNA (Mohun *et al.*, 1984) and methods described by Maniatis *et al.* (1982). The location and orientation of the gene were established using subcloned regions of the cDNA as hybridization probes. Restriction fragments containing the entire gene and flanking DNA were subcloned into pUC12. DNA sequencing was performed using the dideoxy method (Sanger *et al.*, 1977) with M13 libraries obtained by the random, 'shotgun' procedure (Bankier and Barrell, 1983). Sequences were compiled using the DB programs of R.Staden (1982, 1984) and analysed using the ANALYSEQ and DIAGON programs. Each nucleotide was sequenced an average of four times on each DNA strand. The gene sequence has been entered into the EMBL Nucleotide Sequence Data Library.

Construction of cardiac actin gene derivatives

A 603-nucleotide fragment containing the cardiac actin gene promoter and cap site was obtained by insertion of an *Xho*I linker into a *Bst*N1 site within the first exon and subsequent *Bal*31 digestion to remove all but 23 nucleotides of the exon from the promoter fragment. After insertion between the *Eco*RI and *Sma*I sites of pUC12, this was re-excised as an *Eco*RI–*Bam*HI fragment. The adjacent 2.5-kb *Kpn*I–*Eco*RI fragment flanking the 5′ end of the promoter was subcloned alongside the modified promoter fragment into the vector pUC18CAT. This contains a *Hind*III–*Bam*HI fragment from pSVoCAT (encompassing the bacterial CAT gene and SV40 polyadenylation site) cloned as a blunt-ended fragment into the *Hinc*II site. The final construct therefore contained a 3-kb fragment of *Xenopus* DNA containing the cardiac actin gene promoter and cap site fused to the bacterial CAT gene. 5′ deletion mutants of this chimaeric gene were constructed using the unique *Eco*RI, *Sca*I and *Bal*I sites within the actin promoter. In each case the unique site was filled in with the Klenow fragment of DNA polymerase and a *Kpn*I linker added. After digestion with *Kpn*I, the truncated actin promoter was subcloned as a *Kpn*I–*Bam*HI fragment into pUC18CAT. The resulting fusion genes contained 580, 416 and 217 nucleotides of actin gene promoter sequence. A further 5′ truncation to −56 (removing the CAAT box) was obtained by removal of a *Sac*I fragment from the shortest deletion mutant.

Microinjection of oocytes and eggs

DNA was injected into eggs or into uncentrifuged oocytes of *X. laevis* as described by Gurdon and Wakefield (1986). Each oocyte germinal vesicle was injected with 20 nl of solution containing from 0.1 to 20 ng of DNA, as indicated in the figures. Oocytes were then cultured for 16–20 h at 19°C before freezing.

Fertilized eggs were injected at the two-cell stage, usually with 10 nl containing 250 pg of DNA into each blastomere at a point between the animal pole and the equator. Eggs were dejellied using cysteine hydrochloride prior to injection and kept in Modified Barth Solution; (MBS) Gurdon and Wakefield, 1986] containing 2% Ficoll (Pharmacia) during injection. They were then cultured until stage 6 (Nieuwkoop and Faber, 1956) in full strength MBS before being transferred to 1/10 MBS and cultured until stage 18. Embryos were stored frozen (−70°C) for subsequent analysis.

Fertilized eggs very often undergo aberrent cleavage and abnormal later development if injected with more than 500 pg of any kind of DNA (Gurdon, 1974) making it difficult to dissect the dorsal region of a stage 18 neurula. The injection of smaller amounts of DNA greatly reduces the amount of gene expression from injected DNA. Conjugates comprising animal and vegetal portions of blastulae after the removal of the equatorial region are readily cultured until control embryos have reached stage 18. Unlike intact embryos, the conjugates are largely unaffected by injection doses as large as 500 pg.

CAT assays

Frozen oocytes, embryos and embryo fragments were homogenized directly in ice-cold 0.25 M Tris-HCl (pH 7.8) using 30–50 μl of buffer per oocyte or embryo. Homogenates were spun at 12 000 r.p.m. for 10 min in the cold and the yolk-free supernatant stored at −70°C. Enzyme assays were performed as described by Gorman *et al.* (1982) and analysed by t.l.c. Levels of actetyl-^{14}C-chloramphenicol synthesized were measured by scintillation counting and values corrected for variations in loading on the t.l.c. plate. Each assay was repeated at least twice.

RNA mapping

Total RNA was isolated from oocytes and embryos as described previously (Mohun *et al.*, 1984). Cardiac actin gene transcripts were mapped by RNase protection as described by Zinn *et al.* (1983). Nuclease-resistant fragments were fractionated on thin 6–8% acrylamide sequencing gels and their sizes estimated from DNA markers or from a DNA sequence electrophoresed in parallel. Plasmids used to prepare RNA probes were constructed by transferring fragments obtained in the random sequencing procedure from M13mp8 into the appropriate orientation Sp6 vector. pSpAlu26 comprises a 267-nucleotide fragment encompassing the first exon of the gene and was used to map the 5′ end of cardiac actin transcripts. pSp9A75 contains a fragment of approximately 500 nucleotides that spans the 3′ most exon of the gene. This was used to map the 3′ termini of transcripts. Probes were synthesized as described by Melton *et al.* (1984).

Acknowledgements

We would like to thank Sharon Fairman for technical assistance and Barbara Rodbard for preparation of the manuscript. Funding from the Cancer Research Campaign is gratefully acknowledged. We also thank C.Wilson and H.R.Woodland for helpful discussions of the microinjection procedure. T.M. is supported by the Research Fellowship in Biological Sciences of the Royal Society.

References

Andres,A.C., Muellener,D.B. and Ryffel,G.U. (1974) *Nucleic Acids Res.*, **12**, 2283–2302.

Bankier,A.T. and Barrell,B.G. (1983) In Flavell,R.A. (ed.), *Techniques in the Life Sciences*. Elsevier Scientific Publishers Vol. B5, pp. 1–34.

Bendig,M.M. (1981) *Nature*, **292**, 65–67.

Bendig,M.M. and Williams,J.G. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6197–6201.

Bendig,M.M, and Williams,J.G. (1984) *Mol. Cell. Biol.*, **4**, 567–570.

Bienz,M. (1984) *EMBO J.*, **3**, 2477–2483.

Bisbee,C.A., Baker,M.A., Wilson,A.C., Hadji-Azimi,I. and Fischberg,M. (1977) *Science*, **195**, 785.

Buckingham,M.E. and Minty,A.J. (1983) In *Eucaryotic Genes: Their Structure, Activity and Regulation; Contractile protein genes*. Butterworths, London, pp. 365–395.

Busby,S.J. and Reeder,R.H. (1983) *Cell*, **34**, 989–996.

Etkin,L.D. and Balcells,S. (1985) *Dev. Biol.*, **108**, 173–178.

Etkin,L.D., Pearman,B., Roberts,M. and Bektesh,S.L. (1984) *Differentiation*, **26**, 194–202.

Ferguson,B., Jones,N., Richter,J. and Rosenberg,M. (1984) *Science*, **224**, 1343–1345.

Gali,G., Hofstetter,H., Stunnenberg,H.G. and Birnstiel,M.L. (1983) *Cell*, **34**, 823–828.

Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.

Grichnik,J.M., Bergsma,D.J. and Schwartz,R.J. (1986) *Nucleic Acids Res.*, **14**, 1683–1701.

Gunning,P., Mohun,T., Ng,S., Ponte,P. and Kedes,L. (1984) *J. Mol. Evol.*, **20**, 202–214.

Gurdon,J.B. (1974) *Nature*, **248**, 772–776.

Gurdon,J.B. and Brown,D.D. (1977) In T'so,P. (ed.), *The Molecular Biology of the Mammalian Genetic Apparatus*. Amsterdam North-Holland Publ. Co., Vol. 2, pp. 111–123.

Gurdon,J.B. and Smith,D. (1984) In *Transfer and Expression of Eukaryotic Genes.*, Academic Press, New York, pp. 143–151.

Gurdon,J.B. and Wakefield,L. (1986) In Celis,J.E., Graessmann,A. and Loyter,A. (eds), *Microinjection and Organelle Transplantation Techniques: Methods and Applications*. Academic Press, London.

Gurdon,J.B., Fairman,S., Mohun,T.J. and Brennan,S. (1985) *Cell*, **41**, 913–922.

Krieg,P.A. and Melton,D.A. (1985) *EMBO J.*, **4**, 3463–3471.

McKnight,S.L. and Kingsbury,R. (1982) *Science*, **217**, 316–324.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.*, Cold Spring Harbor Laboratory Press. NY.

Mattaj,I.W., Lienhard,S., Jiricny,J. and De Robertis,E.M. (1985) *Nature*, **316**, 163–167.

Melloul,D., Aloni,B., Calvo,J., Yaffe,D. and Nudel,U. (1984) *EMBO J.*, **3**, 983–990.

Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.

Minty,A. and Kedes,L. (1986) *Mol. Cell. Biol.*, **6**, 2125–2136.

Mohun,T.J., Brennan,S., Dathan,N., Fairman,S. and Gurdon,J.B. (1984) *Nature*, **311**, 716–721.

Nakamura,O., Takasaki,H. and Ishihara,M. (1970) *Proc. Jap. Acad.*, **47**, 313–318.

Newport,J. and Kirschner,M. (1982) *Cell*, **30**, 687–696.

Nieuwkoop,P.D. and Faber,J. (1956) *Norman Table of Xenopus laevis (Daudin)*, Amsterdam North-Holland.

Nudel,U., Greenberg,D., Ordahl,C.P., Saxel,O., Neumann,S. and Yaffe,D. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3106–3109.

Ordahl,C.P. and Cooper,T.A. (1983) *Nature*, **303**, 348–349.

Rusconi,S. and Schaffner,W. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 5051–5055.

Sanger,F., Nicklen,S. and Coulsen,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.

Singh,H., Sen,R., Baltimore,D. and Sharp,P.A. (1986) *Nature*, **319**, 154–158.

Slack,J.M.W. (1983) *From Egg to Embryo*. Cambridge University Press, Cambridge, UK.

Staden,R. (1982) *Nucleic Acids Res.*, **10**, 4731–4751.

Staden,R. (1984) *Nucleic Acids Res.*, **12**, 521–538.

Sturgess,E.A., Ballantine,J.E.M., Woodland,H.R., Mohun,P.R., Lang,C.D. and Dimitriadis,G.J. (1980) *J. Embryol. Exp. Morph.*, **58**, 303–320.

Stutz,F. and Spohr,G. (1986) *J. Mol. Biol.*, **183**, 349–361.

Sudarwati,S. and Nieuwkoop,P.D. (1971) *Wilhelm Roux' Arch.*, **166**, 189–204.

Wilson,C., Cross,G.S. and Woodland,H.R. (1986) *Cell*, in press.

Zakut,R., Shani,M., Givel,D., Newman,S., Yaffe,D. and Nudel,U. (1982) *Nature*, **298**, 857–859.

Zin,,K., DiMaio,D. and Maniatis,T. (1983) *Cell*, **34**, 865–879.

Received 5 September 1986